Implementing a Secure Annotation Service

Imran Khan, Ronald Schroeter, Jane Hunter
Annotation Systems

- **Purpose:**
  - “Value-adding”/interpretation of digital objects/scientific data
  - Secondary metadata, semantic tagging
  - Facilitate group discourse and harness collective knowledge
  - Share amongst trusted colleagues distributed across institutions
  - Share across distributed annotation servers

- **Usage:**
  - Education
  - eResearch/eScience Collaboratories
  - Ethnographic Media Analysis
  - Telemedicine
Existing Clients

- Amaya (Client) + Annotea (Server)
- Annozilla
- Photo Annotation Systems
  - Mindswap
- Proprietary systems
  - Word (comments)
  - PDF (comments)
Limitations of Current Systems

- Don’t support search and retrieval across distributed servers
- Lack of security/privacy mechanisms
- Lack of responsiveness
  - Slow, no notification method
- Limited search capabilities
- Limited media types supported
- Coarse granularity
- Unstructured annotations
  - Flat text only
Proposed solution

- 2 Levels of Protection:
  - Authentication to annotation server
    - login id / password
  - Access constraints over annotations
    - create, list, read, edit, delete

- Combine 3 Core Technologies:
  - Annotea
  - Shibboleth
  - XACML
Shibboleth

- Federated Identity Management
- Single Sign On
- Separates authentication (IDP) and authorization (SP)
- Access control definition not dynamic/extendible
Shibboleth & Annotea Server

- Shibboleth provides assertions (user attributes)
- XACML module (Sun API) makes access decisions
- Annotation Server sits behind to handle requests
XACML (request + policy = decision)

Request

- **Subject**
  - `eduPersonAffiliation = "student"`

- **Action**
  - `Action-id = "read"`

- **Resource**
  - `Resource-id = "https://server/AnnotesServlet/body/123"`

My Policy

- **Staff (eduPersonAffiliation = "staff")**
  - list = "Permit"
  - read = "Permit"
  - read policy = "Deny"

- **Student (eduPersonAffiliation = "student")**
  - list = "Permit"
  - read = "Deny"
  - read policy = "Deny"

-> access denied / granted
Annotea

- RDF-based
  - Open
  - Extensible
  - Semantic Web
- Existing Servers:
  - Zope
  - PerlLib

[Ref: http://www.w3.org/2001/Annotea/Papers/KCAP01]
Annotea Extensions

- Multimedia
- SVG drawings
- XACML policies
- Any other domain specific information

[Ref: http://www.w3.org/2001/Annotea/Papers/KCAP01]
Demo

- Demo video of annotation clients usage
Vannotea – Collaborative Multimedia Annotation
Future Work

- Thorough user evaluation, usability testing
  - Ethnography, Crystallography, Indigenous Knowledge
- Reduced reliance on Shibboleth
- Extending Annotea further: Ontology-based
- Annotation of:
  - links between multiple objects, PDF files, Excel
- SPARQL querying
- Overview of annotations across collections
- RQF – ranking, peer review
- Post-processing of annotations
Deliverables

- Developed 2 Annotea clients
  - Browser plug-in
  - Vannotea – collaborative multimedia
- Single-sign on to distributed annotation servers
- Secure access and controlled sharing
- Source is authenticated -> trust
- Can annotate text, HTML, images, video, audio, 3D
  - Fine-grained regions/segments
- Structured annotations
  - ranking, comments, related resources, keywords
- Support textual, URLs or files
- Search, browse and retrieval
  - across metadata (creator, date, language) and content
- Staggered, threaded – discussions
- RSS feeds – notify of changes, responses